DOCUMENT RESUME

ED 263 780 FL 015 402

AUTHOR German, Patricia; Perkins, Kyle
TITLE Cloze Residual Gain: A Procedure for Measuring Information Gain in ESL Reading Comprehension.
PUB DATE 84
NOTE 11p.; In: Language Key to Learning. Selected Papers from the Annual State Convention of the Illinois Teachers of English to Speakers of Other Languages/Bilingual Education (12th, 1984). Volume V. (FL 015 399).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Gains; *Cloze Procedure; College Students; Educational Research; *English (Second Language); Foreign Students; Higher Education; *Measurement Techniques; *Reading Comprehension; Research Methodology; Second Language Instruction; Statistical Analysis; *Test Interpretation

ABSTRACT

This study used the residual gain statistic in reading comprehension research in English as a second language to assess individual differences in improvement accruing from learning and training. The subjects were 34 foreign college students with a variety of native languages. A 35-item cloze test with every eighth word eliminated, but with the first and last sentences unmutilated, was administered in two linguistics class sections. A week later, students were given the passage intact and allowed to study it. The test was then readministered. A standard reading comprehension test was used as a criterion measure. Half the subjects' performance was higher than predicted on the basis of pretest scores, and half the subjects' performance was lower than predicted, an unanticipated result. However, the technique did allow a partitioning of prior knowledge from that acquired by reading, and the results identified superior and inferior "improvers" for further research on the speed and quality of learning in reading. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Cloze Residual Gain: A Procedure for Measuring Information Gain in ESL Reading Comprehension

Patricia German and Kyle Perkins

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Richard A. Orem

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED263780

Abstract

This paper reports the use of the residual gain statistic in ESL reading comprehension research. The residual gain statistic is used to assess individual differences in improvement which accrue from learning and training. Subjects are equated on the basis of the pre-training measure, and the influence of regression effects are reduced.

In this study half the subjects scored higher than predicted on a cloze test after reading the unmutilated passage. The other half scored lower than predicted. The predictions were based on cloze scores before the subjects read the intact passage. The analysis identified a pool of superior improvers and a pool of inferior improvers.

1. Introduction

Previous research has shown that in testing reading comprehension it is necessary to distinguish between general language competence and/or common world knowledge, and information gleaned from the passage by the reader during the reading process. The usual paradigm for this type of research involves administering a test, providing the subjects with some sort of training, and, finally, administering the test again. Subtracting the first score from the second would give a measure of information gain. According to Harrison and Dolan (1979:19), "it seems that to a certain extent the effects of overall language competence are canceled out, and that IG (information gain) scores offer a truer index of the extent to which the reader has assimilated new information."

The notion of gain scores is a radical departure from the usual methods of assessing reading comprehension, whereby a test is normally administered after the subjects have read the test passage. The importance of and need for gain scores have become increasingly more apparent due to the growing body of research literature on the differing amounts of pre-knowledge about the test passage the subjects bring to the testing hall, and due to the fact that the data from a comprehension test administered after reading reflect both the knowledge possessed before reading the test passage, plus the information gleaned while reading. If we are to get a clear idea of what our reading tests measure, we need some approach to separate prior knowledge from acquired knowledge during the reading process.

Traditionally, gain scores have been analyzed by two different approaches: crude gain and residual gain. Determining crude gain

FL015402

BEST COPY AVAILABLE

BEST COPY AVAILABLE

involves subtracting the pre-test scores from the post-test scores, and the difference between the two is called crude gain. There are at least two major problems with crude gain. First, subjects with comparatively lower pre-test scores tend to make higher gains as measured by post-tests than subjects with average and above-average pre-test scores, and subjects with comparatively higher pre-test scores tend to make lower gains on the post-tests--a sort of plateau or "ceiling" effect. Second, reading research has shown that negative correlations obtain between initial and final status in reading improvement when measured by crude gain (Bloomer 1962; Chansky and Bregman 1967; Kamman 1963; Ramsey 1960; Ranson 1955; and Schneyer 1963).

The difficulties with crude gain have been met with the residual gain statistic developed by Manning and DuBois (1962). Residual gain studies are used to assess individual differences in improvement which accrue from learning and training. The procedure statistically equates subjects on the basis of pre-learning or pre-training measurement, and it does not require that the pre- and post-measures be expressed in equal interval scales. In addition, the influence of regression effects upon the measurement of improvement are reduced.

As previously mentioned, research has indicated that negative correlations may obtain between initial status and crude gain in reading studies. Manning and DuBois (1962:290) believe that such spuriously negative coefficients can be overcome by using residual statistics.

> These difficulties, encountered in measuring change by the absolute difference between two test scores (crude gain), would probably be overcome if there were assurance that increments in scores on both initial and final tests were equivalent, that the two tests have the same zero point, and that the tests are valid and reliable at all levels of proficiency.

The purpose of this study was to apply the residual gain technique to reading gains measured by the cloze procedure in an ESL context. Rankin and Dale (1969:19) provide the following explanation of residual gain:

> Residual gains may be defined as the deviation of a subject's post-test score from the post-test score that was predicted from his pre-test score, or the deviation of scores from the regression line of final on initial scores. The residual gain formula is expressed as:
>
> $$\text{Residual gain} = z(2) - r(12)\, z(1)$$
>
> where $z(2)$ is the post-score in z-score form (z-scores or normalized standard scores equal the number of standard deviations the raw score is found away from the mean score), $r(12)$ is the correlation between pre- and post-tests, and $z(1)$ is the pre-test score in z-score form. By the use of this technique, subjects are statistically equated with respect to initial status on the pre-test, and the correlation between

BEST COPY AVAILABLE

residual gain and initial status is zero.

The following research hypotheses were tested:

1. There is a significant difference in pre- and post-test means.
2. There is a significant correlation between crude and residual gains.
3. There is a negative correlation between pre-test initial status and crude gain.
4. There is no significant correlation between crude gains and the criterion measure.
5. There is no significant correlation between residual gains and the criterion measure.

2. Methodology

2.1 Subjects . This research was conducted in the Department of Linguistics at Southern Illinois University. Thirty-four foreign students enrolled in Linguistics 101, Basic Composition for Foreign Students, served as unpaid volunteers fro the project. Native languages included Chinese, Malay, Spanish, Greek, Arabic, Thai, Yoruba, Swedish, Urdu, Brunei, and Korean.

2.2 Materials . We constructed a thirty-five item cloze test by deleting every eighth word. The passage was 348 words long, and the topic was Finland's form of government. The first and last sentences were left unmutilated. Exact-word scoring was employed.

The cloze test was administered to two sections of Linguistics 101. Working time was forty minutes. One week later, the subjects were given an intact version of the entire passage to read and study for twenty minutes. The passages were collected, and fresh copy of the same cloze test was administered the second time; forty minutes was allotted for the test. Exact word scoring was also used to mark the post-test. No indication was given during the pre-testing that the subjects would be tested again. The California Achievement Test Reading Comprehension section (CAT-R) was used as a criterion measure.

3. Analysis and Results

The descriptive statistics are presented in Table 1. The observed reliability coefficients were somewhat lower than we like for research of this nature. One heartening note is that the reliability of the post-cloze (.77) was higher than the reliability of the pre-cloze (.53). This finding runs counter to what one would expect because gain measures entail using a test twice and therefore there are two opportunities for measurement error to creep in. We believe that the subjects' reading the unmutilated passage and familiarization with the cloze procedure better enabled them to utilize the inherent redundancy in the text, enabling them to analyze by synthesizing better the second time around. The result is that the post-test data more truly exhibit the diversity of the subjects' processing ability, hence better person separability or reliability.

Correlation coefficients for the cloze and criterion measure are reported in Tables 2 and 3 (corrected for attenuation). Again we were surprised by the low magnitude of correlation between the cloze and the reading comprehension test, because cloze is usually a robust measure of reading comprehension.

Hypothesis 1 was tested by submitting the pre- and post-test data to a t-test for dependent samples. The t-value of 11.77 is significant at the .01 level for 33 degrees of freedom and demonstrates that the cloze procedure is a highly sensitive measure of reading gains.

We computed the cloze residual gain scores following Tracy and Rankin (1967:364):

1. Convert both pre- and post-reading test scores to z-scores for each student.
2. Compute the r (correlation coefficient) between pre- and post-test raw scores.
3. Obtain predicted post-test z-scores by multiplying the correlation coefficient by the pre-test z-score for each student.
4. Subtract the predicted post-test z-score from the obtained post-test z-score for each student.

In addition, a second set of residual gain scores was computed based on the predicted regressions of the post-test raw scores on the pre-test raw scores (Taschow 1969). The results are presented in Table 4.

With the z-score residual gain statistics, the pre- and post-test scores are expressed in comparable units; it is therefore easy to compare residual gain with crude gain for particular subjects. Let's consider the six subjects who scored 10 on the pre-test, i.e., subjects 9, 12, 17, 18, 22, and 24 and the two subjects who scored 20 on the pre-test, i.e., subjects 3 and 29.

The six subjects who scored 10 on the pre-test had the pre z-score -.805 and were predicted to have the post z-score -.282. The predicted regression post-score was 21.465. Subject 9 had a post z-score of -.562 (a residual gain of -.280) and a crude post score of 20 (a residual gain of -1.465). Subject 12 had a post z-score of +.017 (a residual gain of +.299) and a crude post score of 23 (a residual gain of +1.535). Subject 17 had a post z-score of +1.369 (a residual gain of +1.651) and a crude post score of 30 (a residual gain of +8.535). Subject 18 had a post z-score of -.176 (a residual gain of +.106) and a crude post score of 22 (a residual gain of +.535). Subject 22 had a post z-score of +1.176 (a residual gain of +1.458) and a crude post score of 29 (a residual gain of +7.535). Subject 24 had a post z-score of -.948 (a residual gain of -.666) and a crude post score of 18 (a residual gain of -3.465).

The two subjects who scored 20 on the pre-test had the pre z-score +2.340 and were predicted to have the post z-score of +.819. The predicted regression post score was 27.118. Subject 3 had a post z-score of +1.176 (a residual gain of +.357) and a crude post score of 29 (a residual gain of +1.882). Subject 29 had a post z-score of -.176 (a residual gain of -.995) and a crude post score of 22 (a residual gain of -5.118).

Taschow (1969:29) offers an explication of the residual gain for the predicted regressions of the post-test raw scores on the pre-test

raw scores--an explanation which also serves by analogy for the z-score residual gain:

> The graphical presentation of the crude gain method is based on y=x which for the purpose of discussion is called the zero line or no-learning line. Progress in student's proficiency is noted by a plus sign before the crude gain number and with its position above the zero line; no learning progress is indicated by ±0 and with its position directly on the zero line; and failure of improvement is marked by a minus sign before the crude gain number and with its position below the zero line.
>
> The graphical method of the residual gain procedure shows the regression line determined by its intercept and its slope. At any given value of x (pre-test), the height of the regression line tells the average value of y (post-test). Based on the linear regression formula y=a+bx, where x is zero, y is equal to a which indicates where the line intercepts the y=axis, and b, the slope or regression coefficient indicates how much y changes with a unit change in x. A student's residual gain in found by locating the interesting lines on the graph for the student's x and y scores. The meeting point of the interesting lines falls on, above, or below the regression line. If the point which represents the student's performance falls directly on the line, the student has performed as predicted. If the point falls above the regression line, the distance between its position and the regression line indicates that the student's performance was higher than predicted on the basis of the pre-test score and a plus sign precedes the residual gain score. If the point falls below the regression line, the distance between its position and the regression line indicates that the student's performance was lower than predicted on the basis of his pre-test scores and a minus sign precedes his residual gain score. Thus, a minus sign does not spell failure of improvement, but tells that the student did not yet improve as much as was expected upon prediction.

Hypothesis 2--there is a significant correlation between crude and residual gains--was confirmed; the obtained correlation coefficient is .96 which is significant at the .05 level.

Hypothesis 3--there is a negative correlation between pre-test initial status and residual gains--was confirmed. The correlation between initial status and the z-score residual gain is -.003, and between initial status and predicted regression residual gain, -.42 which is significant at the .05 level.

Hypothesis 4--there is no significant correlation between crude gains and the criterion measure--was confirmed; the obtained

coefficient is .13.

Hypothesis 5--there is no significant correlation between residual gains and the criterion measure--was confirmed; the obtained coefficient was .19.

We did not expect high magnitudes of correlation between the gain scores and the criterion measure because they measure different constructs. The gain scores relate to the amount of information gleaned during the reading process, and the criterion measure tests identification of main ideas, supporting details, inferencing, and judgment.

The purpose of this research was to utilize the residual gain statistic in ESL reading research. It permits us to measure precisely the differences in improvement of those who have received the same experimental treatment and who have been equated statistically on the basis of the pre-test instrument. Correlates of change in proficiency are important in the evaluation of reading pedagogy and the second language learning theory. An important characteristic of this procedure is that the experimenter can partition prior knowledge from that knowledge acquired by reading. In this particular study half of the subjects' performance was higher than predicted on the basis of pre-test scores and half of the subjects' performance was lower than predicted. Thus, the sample is divided into those who need remediation and those who are ready to move on to other skills. In addition, a pool of superior "improvers" and pool of inferior "improvers" have been identified for research on the speed and quality of learning in reading.

TABLE 1

Descriptive Statistics

Test	No. of Subjects	No. of Items	Mean	SD	SEM	KR-20 r_{tt}	100-Item r_{tt}
Pre-Cloze	34	35	12.56	3.18	2.18	.53	.76
Post-Cloze	34	35	22.91	5.18	2.48	.77	.91
CAT-R	34	44	33.03	5.23	2.51	.77	.88

Post-Cloze, Pre-Cloze $t = 11.77$, df 33, $p < .01$

TABLE 2

Correlation Matrix

	1.	2.	3.
1. Pre-Cloze	1.00		
2. Post-Cloze	.35*	1.00	
3. CAT-R	.22	.26	1.00

* $p < .05$, df 32, two-tailed test

TABLE 3

Correlation Matrix
Corrected for Attenuation

	1.	2.	3.
1. Pre-Cloze	1.00		
2. Post-Cloze	.54*	1.00	
3. CAT-R	.34*	.34*	1.00

*$p < .05$, df 32, two-tailed test

Table 4

The z-Score Method for Obtaining Residual Gains

Student	X Pre-Raw Score	z_x Pre z	Y Post-Raw Score	z_y Post z	$z_{y'}$ Predicted Post z	$z_y - z_{y'}$ Residual Gain	Crude Gain	Regression of Post- on Pre-test: Predicted	Regression of Post- on Pre-test: Residual Gain
1	11	- .491	15	-1.527	-.172	-1.355	4	22.031	-7.031
2	8	-1.434	16	-1.333	-.502	- .381	8	20.335	-4.335
3	20	+2.340	29	+1.176	+.819	+ .357	9	27.118	+1.882
4	14	+ .453	20	- .562	+.159	- .721	6	23.726	-3.726
5	11	- .491	24	+ .210	-.172	+ .382	13	22.031	+1.969
6	17	+1.396	23	+ .017	+.489	- .472	6	25.422	-2.422
7	13	+ .138	14	-1.720	+.048	-1.768	1	23.161	-9.161
8	13	+ .138	16	-1.333	+.048	-1.381	3	23.161	-7.151
9	10	- .805	20	- .562	-.282	- .280	10	21.465	-1.465
10	9	-1.119	18	- .948	-.392	- .556	9	20.900	-2.900
11	14	+ .453	21	- .369	+.159	- .528	7	23.726	-2.726
12	10	- .805	23	+ .017	-.282	+ .299	13	21.465	+1.535
13	12	- .176	19	- .755	-.062	- .693	7	22.596	-3.596
14	11	- .491	25	+ .403	-.172	+ .575	14	22.031	+2.969
15	16	+1.082	20	- .562	+.379	- .941	4	24.857	-4.857
16	14	+ .453	30	+1.369	+.159	+1.210	16	23.726	+6.274
17	10	- .805	30	+1.369	-.282	+1.651	20	21.465	+8.535
18	10	- .805	22	- .176	-.282	+ .106	12	21.465	+ .535
19	14	+ .453	28	+ .983	+.159	+ .824	14	23.726	+4.274
20	13	+ .138	28	+ .983	+.048	+ .935	15	23.161	+4.839
21	12	- .176	20	- .562	-.062	- .500	8	22.596	-2.596
22	10	- .805	20	+1.176	-.282	+1.458	19	21.465	+7.535
23	11	- .491	24	+ .210	-.172	+ .382	13	22.031	+1.969
24	10	- .805	18	- .948	-.282	- .666	8	21.465	-3.465
25	8	-1.434	16	-1.333	-.502	- .831	8	20.335	-4.335
26	13	+ .138	29	+1.176	+.048	+1.128	16	23.161	+5.839
27	7	-1.748	15	-1.527	-.612	- .915	8	19.769	-4.769
28	15	+ .767	23	+ .017	+.268	- .251	8	24.292	-1.292
29	20	+2.340	22	- .176	-.819	- .995	2	27.118	-5.118
30	19	+2.025	29	+1.176	+.709	+ .467	10	26.553	+2.447
31	15	+ .767	28	+ .983	+.268	+ .715	13	24.292	+3.708
32	11	- .491	33	+1.948	-.172	+2.120	22	22.031	+10.969
33	12	- .176	28	+ .983	-.062	+1.045	16	22.596	+5.404
34	14	+ .453	24	+ .210	+.159	+ .051	10	23.726	+ .274

References

Bloomer, R. H. 1962. The Effects of a College Program On a Random Sample of Education Freshmen. Journal of Developmental Reading 5, 110-118.

Chansky, N. M. and B. Bregman. 1957. Improving in Reading in College. Journal of Educational Research 51, 310-317.

Harrison, C. and T. Dolan 1979. Reading Comprehension--A Psychological Viewpoint. In Reading in a Second Language , eds. R. Mackay, B. Barkman, and R. Jordan, 13-23. Rowley, MA:Newbury House.

Kamman, R. A. 1963. Aptitude, Study Habits, and Reading Improvement. Journal of Developmental Reading 6:77-86.

Manning, W. H. and P. H. DuBois. 1962. Correlational Methods in Research on Human Learning. Perceptual and Motor Skills 15:287-321.

Ramsey, W. 1960. An Analysis of Variables Predictive of Reading Growth. Journal of Developmental Reading 3, 158-164.

Rankin, E. F., Jr. and L. H. Dale. 1969. Cloze Residual Gain--A Technique for Measuring Learning Through Reading. In The Psychology of Reading Behavior. Eighteenth Yearbook of the National Reading Conference eds., G. B. Schick and M. M. May, 17-26. Milwaukee: The National Reading Conference, Inc.

Ranson, M. K. 1955. Evaluation of Certain Aspects of the Reading and Study Program at the University of Missouri. Journal of Educational Research 48, 443-445.

Schneyer, J. W. 1963. Factors Associated With the Progress of Students Enrolled in a College Reading Program. Journal of Educational Research 56, 340-345.

Taschow, H. G. 1969. A Comparison of Individual Reading Improvement Scores on a Group of Community College Freshmen as Measured by the Crude Gain Method and the Residual Gain Method. In Psychology of Reading Behavior. Eighteenth Yearbook of the National Reading Conference eds., G. B. Schick and M. M. May, 27-37. Milwaukee: The National Reading Conference, Inc.

Tracy, R. J. and E. F. Rankin, Jr. 1967. Methods of Computing and Evaluating Residual Gain Scores in the Reading Program. Journal of Reading 10, 363-371.